机器学习在提高非自杀性自伤预测力中的应用：一项系统综述*

高白雪 [1] 谢云龙 [1] 罗俊龙 [1 2] 贺 雯 [1 2]

([1] 上海师范大学心理学院，上海 200234)

([2] 上海师范大学教育部“教育大数据与教育决策”实验室，上海 200234)

摘要 非自杀性自伤(NSSI)是重大的公共卫生问题，具有高度污名化、高复杂性和高异质性的特点，传统 NSSI 研究测量和分析方法有限，获得的影响因子预测力较低。近年来机器学习逐步应用于 NSSI 的分析和建模中，并通过提高 NSSI 研究工具预测力、增加预测模型复杂度和精确度、区分 NSSI 类别和亚型，使整体预测性能上升到中等水平。未来需结合 NSSI 传统理论和研究方法使筛选标准更严格，拓展非问卷 NSSI 数据与深度学习、无监督学习的结合，根据“先分型、再迁移”的原则增加模型的可复制性与可比性，进一步提高预测性能。

关键词 机器学习，非自杀性自伤，预测力，应用

1 前言

非自杀性自伤(Non-suicidal Self-injury, 简称 NSSI)指的是个体故意地改变或伤害自己的身体组织，但无死亡意愿(江光荣 等, 2011)。NSSI 是高发性心理障碍，在人群中广泛存在，尤其以青少年群体最为严重，我国青少年自伤检出率在 24.9%～29%左右(Qu et al., 2023; Tang et al., 2018)。NSSI 具有严重的破坏性，不仅造成直接的身体损伤，还是自杀的重要风险指标，甚至比自杀未遂史更能预测自杀行为(Ribeiro et al., 2016)。NSSI 常与其它

收稿日期: 2024-11-29

*上海市哲学社会科学规划项目(项目编号：2022BSH002)资助

通信作者: 贺雯, E-mail: hewen@shnu.edu.cn

精神疾病共病，如抑郁障碍、焦虑障碍、边缘型人格障碍和物质成瘾等(Nitkowski & Petermann, 2010; 孙蒙 等, 2020)。因 NSSI 导致的治疗费用和生产力损失等也给个人和政府带来巨大财政负担(Kinchin et al., 2017)。因此，提高对 NSSI 的预测力有利于准确、及时地干预 NSSI 患者，对个体和社会都具有重大价值。

NSSI 的污名化极大影响研究的预测力。结合近几年关于 NSSI 的综述和 Meta 分析发现，问卷调查法已成为 NSSI 研究主流(De Luca et al., 2023; Peel - Wainwright et al., 2021; Xiao et al., 2022)，然而 NSSI 的污名化常使得问卷调查难以推进，或只得到符合社会期望的结果。NSSI 严重的污名化包括患者本身、他人和大众媒体等对 NSSI 的不解、偏见和歧视，NSSI 患者因此产生内疚、羞耻和尴尬感(Fu et al., 2020; Staniland et al., 2021)，导致其在问卷调查中选择隐瞒(Farajzadeh & Sadeghzadeh, 2023)。与此同时，青少年 NSSI 患病率远高于其它群体，常作为研究首选对象，大规模采集青少年的 NSSI 数据不仅要征求本人同意，还要取得教师和父母的同意。NSSI 在青少年群体中存在明显传染效应，青少年意识到朋友自伤可能产生诱发作用(Syed et al., 2020)，也会引起父母和教师的警觉和拒绝。运用实验法研究 NSSI 需模拟自伤疼痛，包含电击、压痛、冷压、热压等任务，更需要考虑伦理问题(颜赟慈, 2015)。因此目前的 NSSI 研究尚缺乏有效的研究工具。

与此同时，由于 NSSI 本身具有高度复杂性，现有研究方法尚不能从全面和系统的角度予以把握。NSSI 影响因素涉及众多方面：生物遗传方面，涉及 FKBP5 基因、5-羟色胺、阿片类系统、多巴胺系统，情绪、控制、疼痛、奖赏相关脑区等异常(白荣 等, 2023; 邓洵 等, 2022)；个人特质方面，涉及情绪调节缺陷、人格障碍、低自尊等；社会因素方面，涉及亲子关系、家庭教育、人际压力等(胡义秋 等, 2023; 蒋家丽 等, 2022; Steinhoff et al., 2021)。单个影响因素没有特别意义，整合生物、社会、心理等众多类型的变量更能提高对 NSSI 发生和发展的预测力(黄任之 等, 2013)。NSSI 产生和发展的理论模型也并不统

一，不同情境和不同个体间要考虑尽可能多的因素(尹慧芳 等, 2022)，极大增加了识别预测和分析的难度。传统分析方法对变量的选择有限，使得只有少数 NSSI 预测因子在分析后保持统计显著性，导致大量信息的丢失(Zhong et al., 2024)；也可能将线性关系强加于更复杂的模型关系，使所建立的预测模型无法完全对应现实生活中的情况(Burke et al., 2019)。NSSI 的元分析发现，在保护因素研究缺乏的情况下(杨佳欣 等, 2023)，NSSI 有关风险因素的预测力极弱(Fox et al., 2015)。

NSSI 研究的高度异质性也是造成预测力低的重要原因。高异质性来自于三个不同的层面：其一是 NSSI 与多种精神障碍共患，在综合医院就诊的自伤患者中，80%以上存在精神障碍，抑郁症、焦虑症和酗酒障碍等 (He et al., 2023; 孙蒙 等, 2020)，使得表现相同的 NSSI 也可能具有不同的风险因素，或与不同精神障碍存在不同的共病模式(He et al., 2023)。其二是 NSSI 本身存在不同亚型，NSSI 的频率、方法、严重程度和功能等方面因人而异(Baer et al., 2020; Daukantaite et al., 2021; Radziwiłłowicz & Lewandowska, 2017)，预测因子在不同年龄、性别等群体间异质性较高，甚至作用方向完全相反(Akbari et al., 2024; Haregu et al., 2023)。目前已有的分型方法主要基于某一维特征，忽略了不同特征之间的复杂组合和相互作用(Wang et al., 2024)。其三是 NSSI 的定义和测量方法不统一，预测因子在不同样本和方法中可能不一致(Rahman et al., 2021)。因此，解决 NSSI 的异质性对提高 NSSI 研究的预测力至关重要。

总之，测量和分析方式有限，NSSI 本身的高污名化、高复杂性和高异质性共同导致了研究的预测力有限。欲提高 NSSI 研究的预测力，需在提高问卷测量效力的基础上，积极寻找其它测量工具。在 NSSI 分析和建模时，应尽可能纳入更多预测因子，并试图筛选更关键的预测因子，建立更复杂、准确的模型。同时，有效区分 NSSI 高度异质性的类别和亚型也是提高预测力的关键。

机器学习(Machine Learning, ML)是一种自动预测建模技术，可以在高维数据中搜索大量解释变量以及识别变量之间的复杂相互作用。目前，机器学习已经广泛应用于阅读障碍、抑郁症等心理问题及障碍的识别和诊断(Bhadra & Kumar, 2022; Castillo-Sánchez et al., 2020; 卜晓鸥 等, 2023; 董健宇 等, 2020)。机器学习在评估 NSSI 发病率、关键影响因子等方面也有了较为成熟的应用(Burke et al., 2019)。机器学习可以分析更多种 NSSI 变量类型，包括文本、图片、视频等类型数据(Lu et al., 2023; Xian et al., 2019)，极大提高了所建模型的预测力；可纳入更多 NSSI 相关变量，对变量间复杂的关联进行建模，产生更强的预测性能(Fox et al., 2019)；能使用随机森林和决策树等可视化 NSSI 预测因子的重要性，构建比传统模型有更高预测能力的决策树模型(Walsh et al., 2017)；以及利用数据驱动方法，根据数据模式来识别 NSSI 亚型，对症状特征的关系和结构进行客观分析(Lanza & Cooper, 2016)，为 NSSI 不同亚型之间的联系提供新的见解(Wang et al., 2024)。本文将在机器学习与 NSSI 最新研究的基础上，对机器学习提高 NSSI 预测力的应用进行阐述，并对当前研究的不足及未来发展方向进行深入思考与展望。

2 方法

对机器学习与 NSSI 的相关研究进行文献搜索，使用的数据库包括知网、维普、万方、百度学术、Web of Science、Elsevier Science Direct、EBSCO 和 PubMed。英文检索关键词“NSSI/Self-injury/Self-harm” AND “Machine Learning/Deep Learning”，中文检索关键词“自伤/NSSI” AND “机器学习/深度学习”。考虑到机器学习应用于 NSSI 领域时间相对较短，研究相对较少，本研究不对发表语言、发表时间、具体机器学习方法和研究对象进行限定。文献纳入标准为：(1)文献为实证类期刊论文和会议论文；(2)文献主题为机器学习用于提高 NSSI 预测力的相关研究；(3)文献使用可验证的研究方法，有翔实的研究结果和客

观的研究结论。依据此标准进行筛选，最后确定纳入本次系统综述的文献数量为 24 篇(见表 1)，图 1 为文献筛选流程。

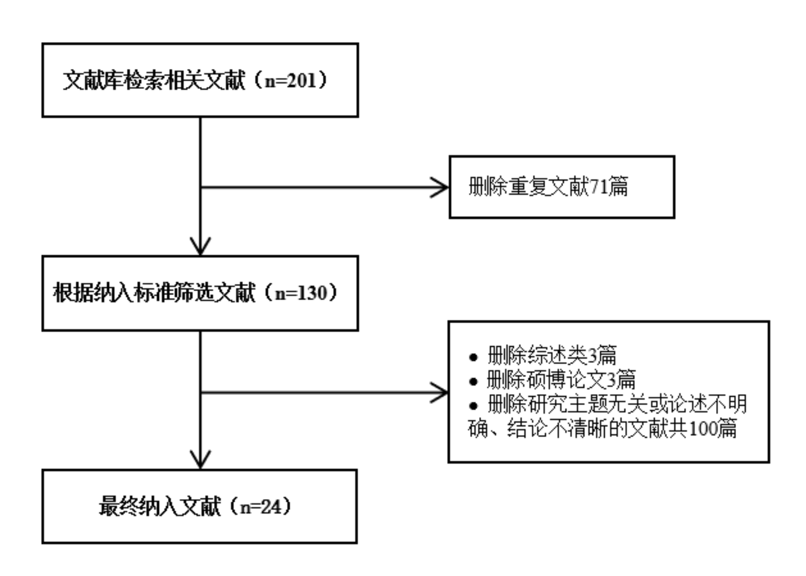


图 1 文献筛选流程图

表 1 机器学习在 NSSI 领域的应用

作者	样本量	纳入变量	分析数据类型	主要研究对象
潘婵 等, 2023	835	19	问卷数据	大学生
Arora et al., 2023	29759	13	人口学数据 问卷数据	自伤成年人
Bao et al., 2024	2385	25	人口学数据 问卷数据	中学生
Chen et al., 2024	114（纵向）	6 个问卷 共 89 项	问卷数据	自伤人群
Cliffe et al., 2021	7188	未说明	临床病例数据 人口学数据	饮食失调患者
Farajzadeh & Sadeghzadeh, 2023	245	19 个问卷 共 662 项	问卷数据	自伤者
Fox et al., 2019	926（纵向）	39	人口学数据 问卷数据	自伤青少年
Guo et al., 2024	5807	21	问卷数据	6-16 岁学生
Kappes et al., 2023	356	512	临床病例数据 问卷数据	精神分裂谱系障碍患者
Kumar et al., 2020	6,037,479	185234	临床病例数据; 人口学数据	自伤患者
Kyron et al., 2021	3690	12	问卷数据	精神病患者
Lang et al., 2024	117（NSSI） 84（正常）	25	人口统计数据; fMRI 数据	自伤者
Lu et al., 2023	87975(结构) 249(非结构)	结构化 4 种 非结构化 4 种	非结构化笔记 人口学数据	囚犯
Marti-Puig et al., 2022	2144	6	活动和情绪数据 EMA 数据	年轻自伤者
Murner-Lavanchy et al., 2024	149（NSSI） 40（正常）	7	临床病例数据	自伤患者
Mürner-Lavanchy et al., 2022	240（NSSI） 49（正常）	24	问卷数据 行为数据 人口学数据	240NSSI 青少年 49 普通青少年
Su et al., 2023	296	497	问卷数据 行为数据 人口学数据	自伤自杀青少年
Swaminathan et al., 2023	721	/	文本信息	有意图自伤者
Xian et al., 2019	1.2 TB 图片、 视频、文本	/	图片数据	自伤者
Xu et al., 2024	112（NSSI） 98（正常）	82	问卷数据 基因数据 人口学数据	自伤者
Yang et al., 2022	186	7	人口学数据 问卷数据	情绪障碍的青少年
Yin et al., 2021	38389	16	临床病例数据	0-17 岁自伤者
Zhong et al., 2024	13304	26	人口学数据 问卷数据	中国西部青少年
Zhou et al., 2024	7967	20	家长及学生 问卷数据	青少年与其父母

3 机器学习在提高 NSSI 预测力的应用步骤

3.1 数据采集

数据采集方法上，机器学习研究首要考虑大数据的需求。只有对大量数据加以训练和测验，机器学习才会产生预测力更高的模型(Yarkoni & Westfall, 2017)。目前有问卷调查、NSSI 患者临床病例和社交媒体数据能满足此要求。由表 1 可见，机器学习在 NSSI 领域采集的数据以问卷为主，NSSI 问卷采集的数据量比其他方法少。数据量最少的为一项纵向研究，间隔六个月重测后共纳入 114 样本(Chen et al., 2024)。值得注意的是，无论是本研究还是 NSSI 相关综述、Meta 分析中，我国大规模问卷调查在样本采集数量和发文数量上都远超其它国家，这体现了我国在大数据研究方面的优势。除问卷收集的数据，回顾式调取患者病例效率更高，样本量更多，普遍包含数万个样本。纳入的文献中最多采集了 6,037,479 份 NSSI 临床病例(Kumar et al., 2020)。相比于问卷数据，临床数据量易采集更多 NSSI 特征，将自然语言处理(Natural Language Processing，NLP)应用于结构化的电子健康记录，可以构建数百个特征用于分析(Arora et al., 2023)。除了问卷和临床病例，社交媒体包含海量 NSSI 相关文字、图片、视频等材料，也满足了机器学习对大样本的需求，一项分析社交平台上自伤流行程度的研究，在 Instagram 采集 1.2TB 材料，包含 170 余万自伤图片和 305K 视频(Xian et al., 2019)，但这些材料量化分析难度更高，因此相关研究极少。

大样本的需求和研究方法的选择制约了对研究对象的选择。问卷调查常以青少年为研究对象，便于采集大量数据，纳入研究中样本量最大的是一项关于我国西部青少年 NSSI 的研究，共采集 13304 样本(Zhong et al., 2024)。除纳入的期刊论文外，我国还有一项针对 NSSI 青少年的大样本研究，共涉及 18 万余中学生、职高生以及大学生(徐小明, 2023)。临床病例除 NSSI 患者外，也会涉及其它精神障碍患者，如进食障碍(Cliffe et al., 2021)、精神分裂谱系障碍(Kappes et al., 2023)等人群。社交媒体数据通常选择 Instagram、Facebook、

Twitter 等热门社交平台，根据 NSSI 相关的话题标签筛选数据(Xian et al., 2019)，但社交媒体的匿名性使研究对象的具体信息无法被确定。

3.2 模型算法选择

收集足够的 NSSI 样本后，需要选择合适的机器学习算法进行分析。从图 2 可观察现有 NSSI 研究中对机器学习算法的选择与使用情况。目前 NSSI 领域的研究以浅层机器学习为主，包括随机森林(Random Forest, RF)、极致梯度提升(Extreme Gradient Boosting, XGBoost)、支持向量机(Support Vector Machines, SVM)等算法。首先是因为 NSSI 研究数据以问卷、人口学等结构化数据为主，而深度学习更擅长从图像或文本等非结构化数据中学习(Kentopp, 2021)。其次，深度学习依赖于更大的数据集以获得更好的性能，当数据稀少且项目间差异不大时，浅层学习性能更优，数据分析也更稳健(Farajzadeh & Sadeghzadeh, 2023)。总之，浅层学习更适用 NSSI 现有数据类型。

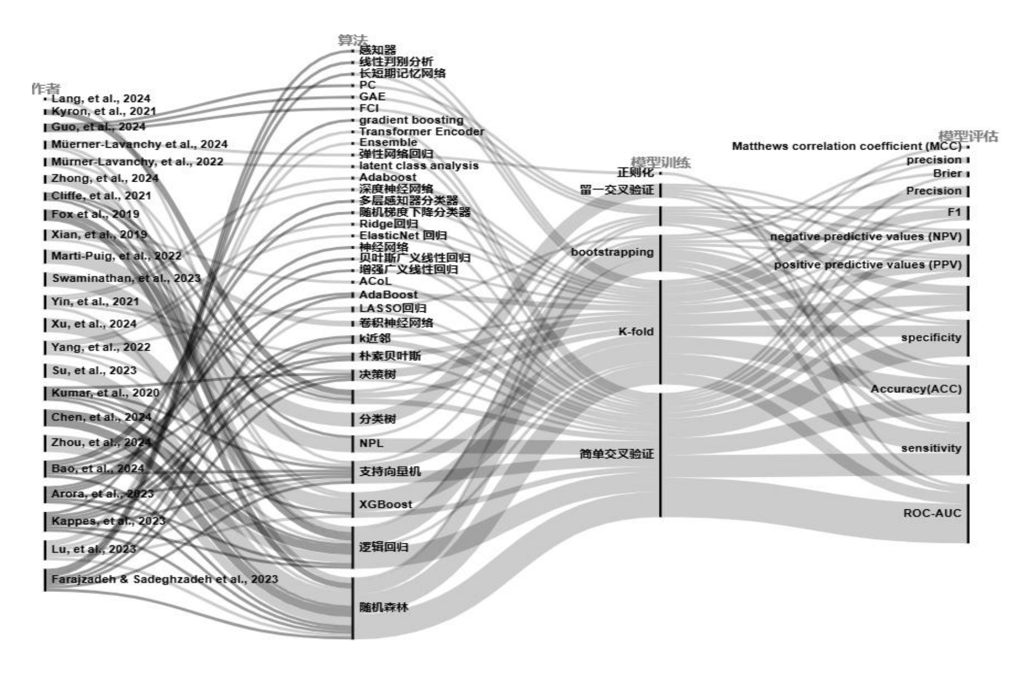


图 2 机器学习在 NSSI 领域应用桑基图

浅层学习中使用频率最高的是随机森林算法。因为 NSSI 是复杂、多因素的，而随机森林算法对 NSSI 数据变化的容忍度高，能有效处理异常数据和分析问题(Jaiswal & Samikannu, 2017)，因此被广泛用于预测自伤行为(Burke et al., 2019)。其次是极致梯度提升

算法，此算法是一个优化的分布式梯度提升库，旨在高效、灵活和便携解决 NSSI 关键特征选择和分类问题(Xu et al., 2024)。另外，随机森林和极致梯度提升算法都包括对 NSSI 预测因子的重要性估计，类似于传统回归模型中对变量重要性的估计(标准化系数等)，因此常与传统逻辑回归相结合，进一步简化模型(Chen et al., 2024; Fox et al., 2019)，即筛选出多种算法中最具影响的核心变量，构建预测力更高的模型(Zhong et al., 2024; Zhou et al., 2024)。

另外，支持向量机算法适用于解决样本量少、非线性、高维的模式识别任务，因此当 NSSI 样本量较少时，支持向量机算法可以避免维数灾难(特征数量过多导致的模型性能下降)和过拟合问题(Yang et al., 2022)，有效选择 NSSI 关键特征，构建更精准的分类模型和预测力模型(Arora et al., 2023; Lang et al., 2024; Lu et al., 2023)。因此支持向量机算法也是 NSSI 领域的应用主流之一。

令人意外的是，尽管无监督学习的应用范围广，功能强大，但在 NSSI 领域中却被忽视。在分析 NSSI 在特征、频率、方法等方面亚型时，潜在类别分析(Latent Class Analysis, LCA)、潜在剖面分析(Latent Profile Analysis, LPA)、K 均值聚类(k-means Clustering)等算法作为典型的无监督学习，已有了成熟的应用(Case et al., 2020; Christoforou et al., 2021; Guerin-Marion et al., 2021; Murner-Lavanchy et al., 2022)。在 NSSI 研究领域，无监督算法却常被忽略其从属于机器学习的重要分支。同时，无监督学习在 NSSI 中应用场景有限，在阅读障碍、抑郁、自杀等其它心理问题中，无监督学习已经广泛应用于关键特征选择和提取(Bhadra & Kumar, 2022; Castillo-Sánchez et al., 2020; 卜晓鸥 等, 2023)，可从复杂原始的数据中(文本、fMRI、EEG、眼动数据等)提取最相关、信息量最大的特征数据，并进行建模分析。但目前机器学习涉及的 NSSI 数据类型较单一，几乎不需要特征提取，因此无监督学习应用有限。

综上所述，NSSI 目前数据更为结构化、样本量相对较少，更适用于浅层学习，尤其是随机森林算法、极致梯度提升算法与支持向量机算法。随机森林算法和极致梯度提升算法还可以对预测因子的重要性进行评估和排序，相比而言，随机森林算法更易于理解和实现，在 NSSI 的应用也最为广泛。极致梯度提升算法需要更多的参数调整和调优，能在复杂的 NSSI 问题上有更高的精度预测和更好的应用前景。另外，无监督机器学习算法在 NSSI 关键特征选择和提取上应用不足，有更大的发展空间。

3.3 模型训练

受限于 NSSI 的复杂性与测量、采集样本量的困难程度，机器学习不但会学习 NSSI 变量间关系，还会学习变量间“噪声”，使建立的模型比真实模型更复杂，即出现“过拟合”现象。为选择 NSSI 最优预测或分类模型，减少误差，则需要对机器学习模型进行训练，常有的方法有交叉验证(留一法、K 折交叉验证、简单交叉验证等)、自举法(bootstrapping)和正则化(董波 等, 2021; 周志华, 2016)。

从图 2 可见，在 NSSI 研究领域，交叉验证中的简单交叉验证是最常使用的模型训练方法，研究者通常简单将 NSSI 数据集随机分为训练集与测试集，训练集通常在总数 60%～80%左右，测试集不参与训练，仅仅用来测试模型效果(Lang et al., 2024; Swaminathan et al., 2023; Zhong et al., 2024)。相比于简单交叉验证，K 折交叉验证能进一步兼顾统计效力(杨柳, 王钰, 2015)，因此在 NSSI 研究中，常使用 5 倍或 10 倍 K 折交叉验证法训练模型——即在 NSSI 原始数据集上进行 5 次或 10 次划分，每次划分都进行机器学习训练和评估，继而将所有结果取平均得到最后的评分(Bao et al., 2024; Kumar et al., 2020; Su et al., 2023; Yin et al., 2021)。

也有研究者强调 NSSI 研究中留一法交叉验证(Leave-One-Subject-Out Cross Validation, LOSO)的重要性。相比于留一法，K 折交叉验证使用全部数据集，基于 NSSI 患者而非

NSSI 特征进行学习，导致模型拟合良好但可能偏离 NSSI 问题本身(Marti-Puig et al., 2022)。另外，大量研究使用了随机森林和极致梯度提升等灵活算法探究 NSSI 问题，此类算法没有可加性或线性假设，并且允许预测因子之间任意交互，更适合使用自举法或重复交叉验证等技术对预测模型进行内部验证，否则训练集会高估模型性能(Collins et al., 2015; Steyerberg et al., 2001)。总之，NSSI 模型训练方法需根据模型样本量和算法共同确定，除涉及随机森林和极致梯度提升算法相关研究使用自举法外，其它研究均以交叉验证法为主，尤其以简单交叉验证的应用最为广泛，常见的划分比例为 70%训练集和 30%测试集，或 80%训练集和 20%测试集等。由于简单交叉验证只进行一次训练和测试，本研究建议将其作为初步评估的手段或样本量过于庞大时的选择。如果 NSSI 样本量不需要节省计算成本，为了获得更可靠的性能估计，应采用更高级的交叉验证方法。在 NSSI 样本量相对少，样本同质性较高的情况下，应考虑留一法交叉验证，可更高效利用数据集。当 NSSI 样本量较大，样本中类别分布不均匀时，应考虑使用 K 折交叉验证以及分层 K 折交叉验证等衍生方法。

3.4 模型评估

如图 2 所示，受试者工作特征曲线下面积(Receiver Operating Characteristic Curve - Area Under Curve, ROC-AUC)是 NSSI 应用中最常用的模型评价指标。AUC 在 0.70 至 0.79 表示模型性能一般，0.80 至 0.89 为良好，0.9 以上为优秀(Simundic, 2009)。纳入文献中 NSSI 模型的 AUC 普遍在 0.74-0.89 之间，整体来说模型性能中良。仅有一项研究模型性能很差，随机森林算法不能基于神经认知特征对 NSSI 和健康人群进行分类(Murner-Lavanchy et al., 2022)。此外，样本越大，模型性能越高，甚至可以达到 0.99(Kumar et al., 2020)；且样本足够大时，深度学习模型的 AUC 比浅层学习表现更优(Yin et al., 2021)；基于自然语言构建 NSSI 模型性能比结构化数据表现更优(Lu et al., 2023; Swaminathan et al., 2023)。

AUC 只是诊断准确度的全局度量，没有报告模型单个参数的任何信息，如敏感性和特异性(Simundic, 2009)。且当参与 NSSI 人数较少时，模型会产生高 AUC，导致几乎没有临床效用(Fox et al., 2019)。故 AUC 应结合其它指标共同评估 NSSI 机器学习的模型，根据图 2，这些指标使用频率依次为灵敏度(Sensitivity)、准确性(Accuracy)、特异性(Specificity)、阳性预测值(Positive Predictive Value, PPV)与阴性预测值(Negative Predictive Value, NPV)、F1、召回率(Recall)等，其中准确性和召回率也采用了 AUC 相同的标准(Franklin et al., 2017)，PPV 与 NPV 多用在 NSSI 的二分类识别和预测中(Chen et al., 2024; Mürner-Lavanchy et al., 2022; Swaminathan et al., 2023; Xu et al., 2024; Zhou et al., 2024)。

4 机器学习在提高自伤预测力的应用与优势

4.1 提高研究工具预测力

机器学习有两条路径应对 NSSI 因测量难而导致预测力低的难题。其一是缩减维度或题数以获得更精简高效的问卷工具；其二是丰富其它有效的测量研究工具。Farajzadeh 等(2022)学者认为，自伤量表问题不宜过多，主题应更明确，还应省略明确询问自伤的问题，可有效避免被试疲劳或说谎。因此使用机器学习将 NSSI 常用的 17 个量表和人口学问卷、测谎题共 662 项题目精简为 22 项，便可准确区分个体是否存在 NSSI 行为(Accuracy = 83.6%)。另外，当机器学习样本量较少时，研究者为了防止过拟合也会精简问卷项目，先使用随机森林算法分析全问卷项目，并持续迭代删除一个最不重要的项目，直到 AUC 不能达到更高为止(Train AUC = 1, 95% CI: 1–1)，最后将问卷的 89 个项目缩减到 34 个(Chen et al., 2024)，获得精简度最高的项目集合。

研究者也尝试基于机器学习方法丰富 NSSI 的研究工具，以提高研究的预测力。一项研究基于经验采样法(Experience Sampling Method, ESM)的研究设计了应用程序，通过三个主要部分收集生态瞬时评估(Ecological momentary assessment，EMA)数据，包括情绪、日

常和自伤活动数据，结合机器学习能有效评估 NSSI 各种情绪状态和功能失调的行为(Accuracy = 84.78%)。此方法采集的数据更丰富、高效、便捷，也使得及时干预 NSSI 行为成为可能(Marti-Puig et al., 2022)。一项正在进行中的研究采用混合方法，通过 EMA 技术采集 NSSI 及情绪评估问卷，每天 3 次，为期 14 天，同时配合可穿戴设备收集参与者的心率、步数和睡眠模式，使用动态结构方程建模和机器学习共同分析 NSSI 内在机制，能更全面地分析并有效预测 NSSI 的想法和行为(Ahn & Lee, 2024)。机器学习也使社交媒体数据很好地应用于 NSSI 的识别和预测，NSSI 患者不太可能向成年人或专业人士寻求帮助，而是更喜欢通过短信或社交媒体或非正式的在线支持平台求助(Belfort & Miller, 2018)，因此会产生巨量关于 NSSI 的图片、视频、评论、动态等，使用机器学习训练分类器区分是否有 NSSI，准确率可达 94%(Xian et al., 2019)，远高于传统研究。总之，精简高效的问卷、丰富的研究工具为提高 NSSI 预测力开辟新道路。

4.2 增加模型复杂度和精确度

NSSI 预测因子众多，具有高复杂性的特点，传统研究可纳入的影响因子不足，导致模型预测力低。NSSI 机器学习研究比传统研究纳入更多的预测因子，多数研究还综合了多类型数据，如人口学数据、行为数据、生理数据等(Lang et al., 2024; Su et al., 2023; Xu et al., 2024)，通过提高模型复杂度以增强对 NSSI 的预测力(Fox et al., 2019)。

考虑到模型的临床适用性，虽然机器学习纳入了更多 NSSI 相关变量，但往往只会保留最重要的预测因子。研究者常利用机器学习给 NSSI 相关变量的重要性排序，进一步筛选出关键预测因子。有研究将支持向量机作为基础学习器，并使用二进制蜻蜓算法来找到最小化目标函数的特征组合，发现性别、青少年教育水平、儿童时期的身体虐待、不接受情绪反应以及偏执、边缘化和表演性人格特征与 NSSI 风险增加有关(Yang et al., 2022)。也有研究在机器学习基础上挑选 15 个因子，再使用逐步回归进一步剔除，最终保留预测力最

大的 11 个因子(Zhong et al., 2024)，或者使用排列重要性为预测变量的重要性排序(Su et al., 2023)，选择前 25 个作为关键预测因素。总之，模型算法不同，变量重要性排序的方式、保留数量也没有统一标准，这也造成各个研究选择出的关键预测因素异质性较高。将已纳入文献的关键变量进行总结，见表 2：

表 2 机器学习筛选出 NSSI 关键预测因子

类型	结果
人口学变量	性别(女性)、年龄、与谁(父亲)一起生活、教育水平、除居住地以外的任何地方的入院来源、未定居的住所、白人种族、年轻和失业；民族；社会经济地位
情绪因素	心理痛苦、抑郁、情绪反应性、自我厌恶、焦虑不安、压力、情绪失控、容忍、避免疼痛；非适应性认知情绪调节策略；管理消极情绪自我效能感；情绪调节自我效能感；表达积极情绪自我效能感
认知因素	抽象思维问题、沉思、一般智力低于正常组；有害思想
家庭因素	家庭、家庭动态因素、家庭功能、家庭冲突和父母抑郁、父母精神困扰的求助行为、家庭团聚；童年遭受身体虐待；父母依恋
人际关系	偏离同伴关系；同伴依恋；积极沟通
自伤史	前一个月的自我切割\未来 NSSI 的自评可能性、自我伤害想法、自杀想法和行为、死亡愿望、自残史
精神障碍因素	物质使用障碍、药物滥用；精神共病，饮食失调患者、边缘型人格障碍；精神病史、网络成瘾、COVID-19 相关 PTSD、学业焦虑和睡眠、心理病理学和药物治疗、心理治疗；精神疾病家族史
生理因素	NSSI 生物学表型，主要由低催产素水平、高白细胞和疼痛敏感性降低组成；NTRK2 基因；兴奋性和抑制性神经元异常；脑部疾病
人格因素	偏执、边缘化和表演性人格
学校因素	学校心理意识；学校参与
其它因素	含糖饮料消费、疫情引起的生活方式改变、屏幕使用时间；中毒；窒息；行为问题
保护性因素	生活满意度、青少年的积极发展；身体健康

在 NSSI 的影响因素中，家庭因素、情绪因素、精神障碍更多作为关键预测因子，这与已有研究一致。家庭因素是 NSSI 生态系统中最重要的预测因子(尹斐 等, 2023)，家庭层面的风险因素具有跨时间稳定性(王玉龙 等, 2024)。NSSI 患者存在情绪调节障碍，通过自伤进行情绪调节和控制是 NSSI 主要目的之一(Hasking et al., 2017, Hasking et al., 2023; Lang et al., 2024)。而精神障碍与 NSSI 有高共患率，也是重要风险因素。

此外，机器学习也为探索 NSSI 预测因子带来了新见解，因为机器学习部分脱离传统心理学先理论假设再验证的模式，尽可能多地纳入相关变量，能发现 NSSI 现有理论外的关键影响因素：如屏幕使用时间的增加和含糖饮料的摄入与 NSSI 风险增加有关，青少年

的积极发展可减轻家庭功能障碍和网络成瘾对 NSSI 的不利影响(Guo et al., 2024)。抽象思维受损作为 NSSI 的关键影响因子也出乎意料，可能是抽象思维方面有困难的人往往在社交和解决问题方面也有困难，导致了 NSSI 行为的出现(Kappes et al., 2023)。这些发现进一步拓展 NSSI 的理论与实证研究，提高了 NSSI 模型预测力。

4.3 区分 NSSI 类别和亚型

NSSI 定义与测量不统一，并且存在高度异质性的类别和亚型，是造成模型预测力低的重要原因。机器学习从两方面努力解决高度异质性带来的困境：一方面是在 NSSI 不同定义和测量方法、不同群体和亚型间寻找共同预测因子，另一方面是试图有效区分 NSSI 不同类别与亚型，在同质群体建立更高预测力模型。因此，机器学习在 NSSI 领域的分类有三大类型：其一是有无 NSSI 的分类，以确定 NSSI 关键预测因子，有研究使用机器学习基于精神病史、犯罪史、精神病理学和药物治疗在内的十个变量的组合进行有无自伤分类，达到 68%准确率，71%AUC 的良好预测能力。另一项研究使用机器学习分析综合临床数据，包括唾液和血液样本、心率变异性和疼痛敏感性，并将 NSSI 和相关精神病理学的生物标志物模式进行建模，发现激素和炎症标志物以及疼痛敏感性的组合能够有效区分是否有 NSSI 障碍(Murner-Lavanchy et al., 2024)，有助于纵向预测临床结果或治疗反应。而目前无法根据神经认知特征进行有无自伤分类，意味着此部分认知神经并不是 NSSI 关键特征(Murner-Lavanchy et al., 2022)。其二是不同 NSSI 类别的区分，目前主要是将 NSSI 与其它类别心理问题或精神障碍进行区分，因为 NSSI 与不同心理问题和精神障碍的高共病率可能掩盖其真实影响因子。一项研究使用 NPL 对 721 条电子病历聊天信息进行了处理和系统的训练(Swaminathan et al., 2023)，最终成功将 NSSI 与自杀/杀人意念、家庭暴力等危机进行分类(AUC = 0.98)；还发现疼痛回避作为一种独特的预先存在的因素，可以区分青少年的自杀和 NSSI，但 NSSI 模型的性能相对于自杀表现更差(Bao et al., 2024)；也有通过机器

学习区分自伤和暴力袭击，发现深度神经网络和自适应增强算法预测更准确(Yin et al., 2021)。其三是使用无监督算法对 NSSI 内部功能、频率、目的等进行分类，即探究 NSSI 亚型，以期望在同型间建立高预测力模型。有研究根据 NSSI 特征和严重程度对青少年抑郁症临床样本进行亚型区分，使用潜在类别分析确定了“高自杀意念 NSSI 组”和“低自杀意念 NSSI 组”两种亚型，并发现女孩、单亲家庭、留守经历、抑郁和同伴欺凌增加了属于“高自杀意念 NSSI 组”的概率(He et al., 2023)。另一项研究使用潜在剖面分析发现 NSSI 的动机存在五种亚型，且随时间推移，动机为强烈渴望与他人交流或惩罚自己的被试会更频繁出现 NSSI 行为(Dixon-Gordon et al., 2022)。有研究者分析 21 篇 NSSI 分型的研究后发现，目前 NSSI 分型结果存在明显的异质性，相关分型的证据也不足，但强调了 NSSI 功能和心理疼痛在 NSSI 分型中的重要作用(Wang et al., 2024)。

有无 NSSI 的分类发掘了更多的关键预测因子(Lang et al., 2024; Xu et al., 2024)，在复杂 NSSI 数据中建立预测力更高模型(Lang et al., 2024)。将 NSSI 与其它心理障碍或精神障碍进行分类不仅有重要的研究价值，还有很高的临床应用价值，有利于临床危机信息分类和分诊，节约诊断时间，并降低诊断失误(Swaminathan et al., 2023; Yin et al., 2021)。

NSSI 亚型的异质性高，则进一步说明了 NSSI 问题本身的复杂性。现有证据表明，心理病理学在许多方面都是异质的，不同的病因机制可能与同一疾病有关，并且在同个人中可能存在多种结果(Feczko et al., 2019)。因此进一步探究 NSSI 的分类问题，不仅有利于明晰 NSSI 本身问题，也为其他心理病理学分类识别提供参考。

总之，机器学习有望解决 NSSI 测量困难、高度复杂性和高度异质性的难题，提高了 NSSI 模型的预测力，如图 3 所示：

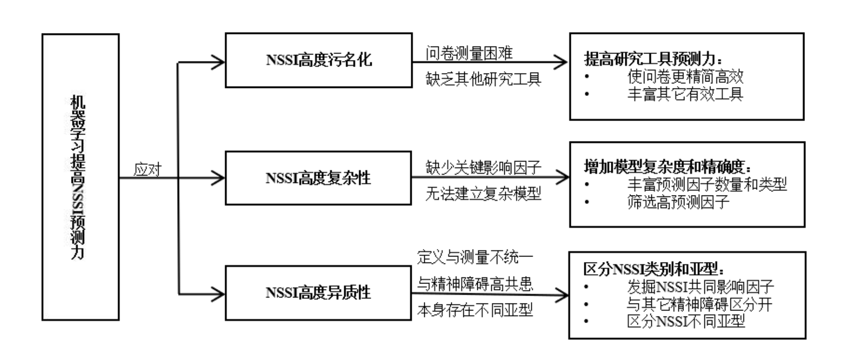


图 3 机器学习提高 NSSI 预测力的应用

5 现存问题与未来展望

5.1 进一步丰富数据类型，以提高模型性能

通过机器学习构建 NSSI 的模型已经达到中良性能，较传统研究已经有了很大的改善。但对比机器学习在其它心理健康/精神问题的应用，NSSI 与机器学习的相关研究显然还有提升空间。NSSI 测量困难导致样本量偏少，过于依赖问卷或临床病例单一类型数据，是影响机器学习性能的关键。未来研究可重点考虑拓展多类型多指标 NSSI 模型和多测量方法进行改善。

在测量方面，生态瞬时评估法(EMA)是适用 NSSI 机器学习最有前景的方法。我国虽然因缺乏统一化的医疗信息管理系统等因素导致病例数据收集困难(董健宇 等, 2020)，但在统筹问卷调查中具有优势，可将问卷与生态瞬时评估技术进一步结合，以实时、真实的方式评估暴露和结果，从而获取大量数据。生态瞬时评估不会重新激活 NSSI 患者的负面想法，结合机器学习分析可准确地分析自伤日常情况及其内外部触发因素(Gee et al., 2020; Martinez-Ales & Keyes, 2019)。

结合多模态多指标数据，也是提高 NSSI 模型性能的重要途径。一项针对囚犯 NSSI 行为的研究发现，NSSI 非结构数据集的模型性能要远高于单一问卷数据集，甚至高于二者结合的综合数据集(Lu et al., 2023)。文本数据比结构化数据包含更多信息，单独使用时产生 AUC-ROC 就达到 0.862，灵敏度、特异性分别达到 0.816 和 0.738。但只使用非结构文本也

带来识别风险，如针对临床病例数据的分类依赖于医生文档，其中语法错误、行话和特殊缩写，都会增加 NLP 错误分类的机会(Cliffe et al., 2021)。因此结合多指标、多模态的大量数据才能真正发挥机器学习的优势，也是 NSSI 模型性能突破的关键。

5.2 结合已有理论与研究方法，拓展深层次研究

机器学习在探究 NSSI 高预测因子、关键特征有重要的发现，但如果完全忽视 NSSI 理论和已有研究，会使机器学习偏离心理学研究，并且想穷尽 NSSI 的影响因素也是不可能的(Siddaway et al., 2020)。因此，机器学习作为新兴的研究方法，应与 NSSI 传统理论和研究方法进一步结合，在已有成果上进一步发展。

研究者使用机器学习建立 NSSI 预测模型时，仅有两项研究纳入和筛选的方式较为科学。一项研究从 NSSI 的家庭系统理论出发，选择 20 个相关预测因子(Zhou et al., 2024)。另一项从大数据角度出发，在 4496 个因子中选择 497 个 NSSI 相关因子进行分析(Su et al., 2023)。未来研究可考虑在前期结合 NSSI 理论和机器学习算法共同筛选变量，从庞大数据库选择强预测因子，再结合逐步回归或结构方程建模，有望进一步提高预测精度(Burke et al., 2019)。如 Nock(2009)的 NSSI 整合模型认为 NSSI 是生物、心理、社会三方面共同作用的结果，涉及遗传、童年创伤、负性情绪、负性认知、压力反应、社交能力等诸多预测因子，机器学习可先从不同方面筛选重要预测因子，再建立模型。或如 Hooley 和 Franklin(2018)提出的 NSSI 获益和阻碍模型，可使用机器学习进一步明确 NSSI 行为的获益和阻碍因子，以及两种相反动力因素的复杂作用，有利于提高对 NSSI 行为的预测。

机器学习也应与 NSSI 传统测量深度融合，如 fMRI、fNRIS、EEG、眼动等方法产生的复杂数据和图像式数据，结合无监督学习和深度学习进行数据降维、关键特征提取和分析，而不仅局限于 NSSI 结构化数据和浅层学习的结合。

此外，机器学习研究应该借鉴 NSSI 传统研究存在的问题，即以横断研究为主(何厚建 等 , 2020)，风险因素探究过多，缺乏保护因素和积极因素的探究(杨佳欣 等., 2023)。机器学习应纳入更多类型变量，并进一步提高 NSSI 纵向水平预测力。

5.3 增强可复制性与可比性，以提高应用价值

目前研究中，NSSI 的关键预测因子异质性较高，这既有 NSSI 本身高度复杂性的原因，也有机器学习使用不同被试群体，不同纳入或筛除变量标准，不同分析算法、训练方法和评估指标的原因，导致同类型研究比较困难(Burke et al., 2019)，临床诊断和干预辅助的应用低。

本研究认为，应遵循“先同再异，先多再少”的原则，提高机器学习在 NSSI 领域的可复制性和可比性。“先同”即先提高 NSSI 同类型群体或同亚型的模型预测力，涉及 NSSI 同类型群体如同性别、同年龄段、同精神障碍的群体，涉及 NSSI 本身的类型如同功能、同自伤频率、同目的等亚型，可先使用潜剖面、潜在类别分析或其它聚类算法来确定分型，再选择样本量较多的群体进行探究。因为 NSSI 患者的日常经历存在可变性(Kyron et al., 2021)，最好在控制样本变异性基础上，缩短随访期，以此提高 NSSI 模型的预测能力(Franklin et al., 2017)。“先多再少”是在提高类型同质、样本较多群体的模型预测力后，再在不同类型，样本较小群体中进行迁移学习。迁移学习是将在数据丰富的源域中实现的学习，再转化为数据稀缺且难以获得的相关目标域的过程。研究者曾在大学生方便样本中预训练 NSSI 的深度神经网络模型，随后在住院患者中进行微调，并与住院患者自身构建模型进行比较，结果深度迁移学习虽没有提高分类精度，但可以缓冲过度拟合(Kentopp, 2021)。因此，若迁移学习后模型性能没有明显改变，则意味着不同群体/类型中存在相同影响因子。若模型性能有大幅提升或下降，则确定为不同质群体，考虑构建新模型，以此

在影响因子中合并不同 NSSI 群体，或不同群体中合并相同预测因子，最终建立稳定分类和同类群体的高预测力模型，为临床诊断和干预决策提供参考。

6 研究结论与贡献

本研究共纳入 24 篇研究，通过桑基图可视化分析了机器学习从收集样本到选择算法、训练模型、评估模型的过程。对机器学习提高 NSSI 预测力的应用和优势进行总结，发现机器学习可精简问卷或丰富其它有效研究工具，有望解决 NSSI 测量困难导致研究预测力低的问题。此外，机器学习还通过发现和识别关键预测因子、建立复杂模型和 NSSI 分类与分型进一步提高了研究的预测力。但模型性能和结果可复制性、可比性有待提高。因此，本研究提出先分型、后迁移的建议，先通过数量多的同类型样本构建较高预测力的机器学习模型，再迁移学习到其它群体，既验证模型的可复制性，也考虑不同群体的异质性，最终实现稳定、高效预测 NSSI。此外，机器学习应与已有 NSSI 的理论、测量方法、分析方法深度融合，并在此基础上进行更深度的研究，进一步提高纵向预测水平。

参考文献

*为纳入系统分析的文献

白荣, 高叶淼, 李金文, 刘霞. (2023). 远近端人际压力与 FKBP5 基因对青少年自伤行为的联合影响: 基于发展的视角. *心理学报*, *55*(9), 1477–1488.

邓洵, 陈宁, 王单单, 赵欢欢, 贺雯. (2022). 自伤行为的神经生理机制及共病障碍比较. *心理科学进展*, *30*(7), 1561–1573.

何厚建, 乐发国, 胡茂荣, 唐金香, 罗若云, 任倩怡, 孙晓岚. (2020). 国外非自杀性自伤研究现状——基于 CiteSpace 可视化分析. *现代预防医学*, *47*(20), 3660–3664.

胡义秋, 曾子豪, 彭丽仪, 王宏才, 刘双金, 杨琴, 方晓义.(2023). 亲子关系和父母教育卷入对青少年抑郁、自伤和自杀意念的影响: 挫败感和人生意义感的作用. *心理学报*, *55*(1), 129–141.

黄任之, 丁立平, 黄敏. (2013). 青少年非自杀性自我伤害行为现状、影响因素及干预. *中国临床心理学杂志*, *21*(6), 965–967.

蒋家丽, 李立言, 李子颖, 雷秀雅, 孟泽龙. (2022). 青少年非自杀性自伤行为持续和停止的预测性因素. *心理科学进展*, *30*(7), 1536–1545.

*潘婵, 刘晓容, 石相孜, 赵文欣, 田萌, 陈思远, 张宛筑. (2023). 基于机器学习构建贵州省大学生非自杀性自伤行为的预测模型. *中国学校卫生*, *44*(8), 1198–1202+1206.

卜晓鸥, 王耀, 杜亚雯, 王沛. (2023). 机器学习在发展性阅读障碍儿童早期筛查中的应用. *心理科学进展*, *31*(11), 2092–2105.

孙蒙, 史战明, 陈登国, 吉航西, 罗丽霞, 张渝雪. (2020). 非自杀性自伤与精神障碍关系研究进展. *国际精神病学杂志*, *47*(1), 11–13+24.

徐小明. (2023). *机器学习方法用于构建青少年自杀自伤行为预警模型研究*[博士学位论文]. 重庆医科大学.

颜赟慈. (2015). *自伤行为中疼痛和见血的情绪调节作用*[硕士学位论文]. 华中师范大学.

杨佳欣, 田于胜, 付熙, 李亚敏. (2023). 青少年非自杀性自伤影响因素的研究进展. *中国临床心理学杂志*, *31*(5), 1145–1149.

杨柳, 王钰. (2015). 泛化误差的各种交叉验证估计方法综述. *计算机应用研究*, *32*(5), 1287–1290+1297.

尹慧芳, 徐浩林, 刘肇瑞, 张一娇, 秦丹妮, 黄悦勤. (2022). 青少年非自杀性自伤行为的理论模型研究. *中国心理卫生杂志*, *36*(8), 707–713.

尹斐, 姜文龙, 周郁秋, 杨金伟, 杨楠. (2023). 生态系统视域下青少年非自杀性自伤行为影响因素的优势分析. *中国健康教育*, *39*(6), 517–522.

江光荣, 于丽霞, 郑莺, 冯玉, 凌霄. (2011). 自伤行为研究:现状、问题与建议. *心理科学进展*, *19*(6), 861–873.

王泉泉, 熊昱可, 刘霞. (2019). 基因-脑-行为视角下的自伤行为产生机制. *心理发展与教育*, *35*(4), 495–503.

王玉龙, 赵婧斐, 蔺秀云. (2024). 家庭风险因素对青少年自伤的累积效应及其性别差异. *心理发展与教育*, *40*(2), 240–247.

董健宇, 韦文棋, 吴珂, 妮娜, 王粲霏, 付莹, 彭歆. (2020). 机器学习在抑郁症领域的应用. *心理科学进展*, *28*(2), 266–274.

董波, 陈艾睿, 张明. (2021). 机器学习在解决过拟合现象中的作用. *心理科学*, *44*(2), 274–281.

周志华. (2016). *机器学习及其应用*. 清华大学出版社.

Ahn, C. Y., & Lee, J. S. (2024). Digital phenotyping for real-time monitoring of nonsuicidal self-injury: Protocol for a prospective observational study. *JMIR Research Protocols*, *13*(1), e53597. Advance online publication. doi:10.2196/53597

Akbari, M., Seydavi, M., Firoozabadi, M. A., & Babaeifard, M. (2024). Distress tolerance and lifetime frequency of non-suicidal self-injury (NSSI): A systematic review and meta-analysis. *Clinical Psychology & Psychotherapy*, *31*(1), e2957.

*Arora, A., Bojko, L., Kumar, S., Lillington, J., Panesar, S., & Petrungaro, B. (2023). Assessment of machine learning algorithms in national data to classify the risk of self-harm among young adults in hospital: A retrospective study. *International Journal of Medical Informatics*, *177*, 105164.

Baer, M. M., Tull, M. T., Forbes, C. N., Richmond, J. R., & Gratz, K. L. (2020). Methods matter: Nonsuicidal self‐injury in the form of cutting is uniquely associated with suicide attempt severity in patients with substance use disorders. *Suicide and Life‐Threatening Behavior*, *50*(2), 397–407.

*Bao, J., Wan, J., Li, H., & Sun, F. (2024). Psychological pain and sociodemographic factors classified suicide attempt and non-suicidal self-injury in adolescents. *Acta Psychologica*, *246*, 104271.

Belfort, E. L., & Miller, L. (2018). Relationship between adolescent suicidality, self-injury, and media habits. *Child and Adolescent Psychiatric Clinics*, *27*(2), 159–169.

Bhadra, S., & Kumar, C. J. (2022). An insight into diagnosis of depression using machine learning techniques: A systematic review. *Current Medical Research and Opinion*, *38*(5), 749–771.

Burke, T. A., Ammerman, B. A., & Jacobucci, R. (2019). The use of machine learning in the study of suicidal and non-suicidal self-injurious thoughts and behaviors: A systematic review. *Journal of Affective Disorders*, *245*, 869–884.

Case, J. A. C., Burke, T. A., Siegel, D. M., Piccirillo, M. L., Alloy, L. B., & Olino, T. (2020). Functions of non-suicidal self-injury in late adolescence: A latent class analysis. *Archives of Suicide Research : Official Journal of the International Academy for Suicide Research*, *24*(SUP2), S165–S186.

Castillo-Sánchez, G., Marques, G., Dorronzoro, E., Rivera-Romero, O., Franco-Martín, M., & De la Torre-Díez, I. (2020). Suicide risk assessment using machine learning and social networks: A scoping review. *Journal of Medical Systems*, *44*(12), 205.

*Chen, S. C., Huang, H. C., Liu, S. I., & Chen, S. H. (2024). Prediction of repeated self-harm in six months: Comparison of traditional psychometrics with random forest algorithm. OMEGA-*Journal of Death and Dying*, *88*(4), 1403–1429.

Christoforou, R., Boyes, M., & Hasking, P. (2021). Emotion profiles of university students engaging in non-suicidal self-injury: Association with functions of self-injury and other mental health concerns. *Psychiatry Research*, *305*, 114253.

*Cliffe, C., Seyedsalehi, A., Vardavoulia, K., Bittar, A., Velupillai, S., Shetty, H., ... & Dutta, R. (2021). Using natural language processing to extract self-harm and suicidality data from a clinical sample of patients with eating disorders: A retrospective cohort study. *BMJ Open*, *11*(12), e053808.

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine*, *162*(1), 55–63.

Daukantaite, D., Lundh, L. G., Wangby-Lundh, M., Clareus, B., Bjarehed, J., Zhou, Y., & Liljedahl, S. I. (2021). What happens to young adults who have engaged in self-injurious behavior as adolescents? A 10-year follow-up. *European Child & Adolescent Psychiatry*, *30*, 475–492.

De Luca, L., Pastore, M., Palladino, B. E., Reime, B., Warth, P., & Menesini, E. (2023). The development of Non-Suicidal Self-Injury (NSSI) during adolescence: A systematic review and Bayesian meta-analysis. *Journal of Affective Disorders*, *339*, 648–659.

Dixon‐Gordon, K. L., Turner, B. J., Haliczer, L. A., Gratz, K. L., Tull, M. T., & Chapman, A. L. (2022). Self‐injury motives: A person‐centered examination. *Suicide and Life‐Threatening Behavior*, *52*(4), 812-827.

*Farajzadeh, N., & Sadeghzadeh, N. (2023). NSSI questionnaires revisited: A data mining approach to shorten the NSSI questionnaires. *PlOS ONE*, *18*(4), e0284588.

Feczko, E., Miranda-Dominguez, O., Marr, M., Graham, A. M., Nigg, J. T., & Fair, D. A. (2019). The heterogeneity problem: Approaches to identify psychiatric subtypes. *Trends in Cognitive Sciences*, *23*(7), 584–601.

Fox, K. R., Franklin, J. C., Ribeiro, J. D., Kleiman, E. M., Bentley, K. H., & Nock, M. K. (2015). Meta-analysis of risk factors for nonsuicidal self-injury. *Clinical Psychology Review*, *42*, 156–167.

*Fox, K. R., Huang, X., Linthicum, K. P., Wang, S. B., Franklin, J. C., & Ribeiro, J. D. (2019). Model complexity improves the prediction of nonsuicidal self-injury. *Journal of Consulting and Clinical Psychology*, *87*(8), 684–692.

Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., ... & Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin*, *143*(2), 187–232.

Fu, X., Yang, J., Liao, X., Ou, J., Li, Y., & Chen, R. (2020). Parents' attitudes toward and experience of non-suicidal self-injury in adolescents: A qualitative study. *Frontiers in Psychiatry*, *11*, 538756.

Guérin-Marion, C., Bureau, J.-F., Lafontaine, M.-F., Gaudreau, P., & Martin, J. (2021). Profiles of Emotion Dysregulation Among University Students Who Self-Injure: Associations with parent–child relationships andnon-suicidal self-injury characteristics. *Journal of Youth and Adolescence*, *50*(4), 767–787.

*Guo, X., Wang, L., Li, Z., Feng, Z., Lu, L., Jiang, L., & Zhao, L. (2024). Factors and pathways of non-suicidal self-injury in children: insights from computational causal analysis. *Frontiers in Public Health*, *12*, 1305746.

Haregu, T., Chen, Q., Arafat, S. M. Y., Cherian, A., & Armstrong, G. (2023). Prevalence, correlates and common methods of non-suicidal self-injury in South Asia: A systematic review. *BMJ Open*, *13*(11), e074776.

Hasking, P., Lewis, S. P., & Tonta, K. (2023). A person-centred conceptualisation of non-suicidal self-injury recovery: A practical guide. *Counselling Psychology Quarterly*, 1–22.

Hasking, P., Whitlock, J., Voon, D., & Rose, A. (2017). A cognitive-emotional model of NSSI: Using emotion regulation and cognitive processes to explain why people self-injure. *Cognition and Emotion*, *31*(8), 1543–1556.

He, H., Hong, L., Jin, W., Xu, Y., Kang, W., Liu, J., ... & Zhao, K. (2023). Heterogeneity of non-suicidal self-injury behavior in adolescents with depression: Latent class analysis. *BMC Psychiatry*, *23*, 1–13.

Hooley, J. M., & Franklin, J. C. (2018). Why do people hurt themselves? A new conceptual model of non-suicidal self-injury. *Clinical Psychological Science*, *6*(3), 428-451.

Jaiswal, J. K., & Samikannu, R. (2017). "Application of Random Forest Algorithm on Feature Subset Selection and Classification and Regression". *World Congress on Computing and Communication Technologies (WCCCT)*, Tiruchirappalli, India. doi: 10.1109/WCCCT.2016.25.

*Kappes, J. R., Huber, D. A., Kirchebner, J., Sonnweber, M., Günther, M. P., & Lau, S. (2023). Self-Harm among forensic psychiatric inpatients with schizophrenia spectrum disorders: An explorative analysis. *International Journal of Offender Therapy and Comparative Criminology*, *67*(4), 352–372.

Kentopp, S. (2021). *Deep transfer learning for prediction of health risk behaviors in adolescent psychiatric patients*. [Unpublished doctoral dissertation]. Colorado State University.

Kinchin, I., Doran, C. M., Hall, W. D., & Meurk, C. (2017). Understanding the true economic impact of self-harming behaviour. *The Lancet Psychiatry*, *4*(12), 900–901.

*Kumar, P., Nestsiarovich, A., Nelson, S. J., Kerner, B., Perkins, D. J., & Lambert, C. G. (2020). Imputation and characterization of uncoded self-harm in major mental illness using machine learning. *Journal of the American Medical Informatics Association*, *27*(1), 136–146.

*Kyron, M. J., Hooke, G. R., & Page, A. C. (2021). Prediction and network modelling of self-harm through daily self-report and history of self-injury. *Psychological Medicine*, *51*(12), 1992–2002.

*Lang, A. N., Zhong, Y., Lei, W., Xiao, Y., Hang, Y., Xie, Y., ... & Wang, C. (2024). Neural mechanism of non-adaptive cognitive emotion regulation in patients with non-suicidal self-injury. *Comprehensive Psychiatry*, *133*, 152487.

Lanza, S. T., & Cooper, B. R. (2016). Latent class analysis for developmental research. *Child Development Perspectives*, *10*(1), 59–64.

*Lu, H., Barrett, A., Pierce, A., Zheng, J., Wang, Y., Chiang, C., & Rakovski, C. (2023). Predicting suicidal and self-injurious events in a correctional setting using AI algorithms on unstructured medical notes and structured data. *Journal of Psychiatric Research*, *160*, 19–27.

Martinez-Ales, G., & Keyes, K. M. (2019). Fatal and non-fatal self-injury in the USA: Critical review of current trends and innovations in prevention. *Current Psychiatry Reports*, *21*, 1–11.

*Marti-Puig, P., Capra, C., Vega, D., Llunas, L., & Solé-Casals, J. (2022). A machine learning approach for predicting non-suicidal self-injury in young adults. *Sensors*, *22*(13), 4790.

*Murner-Lavanchy, I., Koenig, J., Lerch, S., van der Venne, P., Hoper, S., Resch, F., & Kaess, M. (2022). Neurocognitive functioning in adolescents with non-suicidal self-injury. *Journal of Affective Disorders*, *311*, 55–62.

*Murner-Lavanchy, I., Koenig, J., Reichl, C., Josi, J., Cavelti, M., & Kaess, M. (2024). The quest for a biological phenotype of adolescent non-suicidal self-injury: A machine-learning approach. *Translational Psychiatry*, *14*(1), 56.

Nock, M. K. (2009). Why do people hurt themselves? New insights into the nature and functions of self-injury. *Current Directions in Psychological Science*, *18*(2), 78-83.

Nitkowski, D., & Petermann, F. (2010). Non-suicidal self-injury and comorbid mental disorders: A review. *Fortschritte der Neurologie-Psychiatrie*, *79*(1), 9–20.

Peel‐Wainwright, K. M., Hartley, S., Boland, A., Rocca, E., Langer, S., & Taylor, P. J. (2021). The interpersonal processes of non‐suicidal self‐injury: A systematic review and meta‐synthesis. *Psychology and Psychotherapy: Theory, Research and Practice, 94*(4), 1059–1082.

Qu, D., Wen, X., Liu, B., Zhang, X., He, Y., Chen, D., ... & Chen, R. (2023). Non-suicidal self-injury in Chinese population: A scoping review of prevalence, method, risk factors and preventive interventions. *The Lancet Regional Health. Western Pacific*, *37*, 100794. https://doi.org/10.1016/j.lanwpc.2023.100794

Radziwiłłowicz, W., & Lewandowska, M. (2017). Deliberate self-injury functions and their clinical correlates among adolescent psychiatric inpatients. *Psychiatria Polska*, *51*(2), 303–322.

Rahman, F., Webb, R. T., & Wittkowski, A. (2021). Risk factors for self-harm repetition in adolescents: A systematic review. *Clinical Psychology Review*, *88*, 102048.

Ribeiro, J. D., Franklin, J. C., Fox, K. R., Bentley, K. H., Kleiman, E. M., Chang, B. P., & Nock, M. K. (2016). Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: A meta-analysis of longitudinal studies. *Psychological Medicine*, *46*(2), 225–236.

Siddaway, A. P., Quinlivan, L., Kapur, N., O'Connor, R. C., & de Beurs, D. (2020). Cautions, concerns, and future directions for using machine learning in relation to mental health problems and clinical and forensic risks: A brief comment on "Model complexity improves the prediction of nonsuicidal self-injury" (Fox et al., 2019). *Journal of Consulting and Clinical Psychology*, *88*(4), 384–387.

Simundic, A.-M. (2009). Measures of diagnostic accuracy: Basic definitions. *EJIFCC*, *19*(4), 203–211.

Staniland, L., Hasking, P., Boyes, M., & Lewis, S. (2021). Stigma and nonsuicidal self-injury: Application of a conceptual framework. *Stigma and Health*, *6*(3), 312–323.

Steinhoff, A., Ribeaud, D., Kupferschmid, S., Raible-Destan, N., Quednow, B. B., Hepp, U., ... & Shanahan, L. (2021). Self-injury from early adolescence to early adulthood: Age-related course, recurrence, and services use in males and females from the community. *European Child & Adolescent Psychiatry*, *30*(6), 937–951.

Steyerberg, E. W., Harrell Jr, F. E., Borsboom, G. J., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. D. F. (2001). Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology*, *54*(8), 774–781.

*Su, R., John, J. R., & Lin, P.-I. (2023). Machine learning-based prediction for self-harm and suicide attempts in adolescents. *Psychiatry Research*, *328*, 115446.

*Swaminathan, A., Lopez, I., Mar, R. A. G., Heist, T., McClintock, T., Caoili, K., ... & Nock, M. K. (2023). Natural language processing system for rapid detection and intervention of mental health crisis chat messages. *Npj Digital Medicine*, *6*(1), 1–9.

Syed, S., Kingsbury, M., Bennett, K., Manion, I., & Colman, I. (2020). Adolescents' knowledge of a peer's non‐suicidal self‐injury and own non‐suicidal self‐injury and suicidality. *Acta Psychiatrica Scandinavica*, *142*(5), 366–373.

Tang, J., Li, G., Chen, B., Huang, Z., Zhang, Y., Chang, H., ... & Yu, Y. (2018). Prevalence of and risk factors for non-suicidal self-injury in rural China: Results from a nationwide survey in China. *Journal of Affective Disorders*, *226*, 188-195.

Walsh, C. G., Ribeiro, J. D., & Franklin, J. C. (2017). Predicting risk of suicide attempts over time through machine learning. *Clinical Psychological Science*, *5*(3), 457–469.

Wang, Z., Li, D., Chen, Y., Tao, Z., Jiang, L., He, X., & Zhang, W. (2024). Understanding the subtypes of non-suicidal self-injury: A new conceptual framework based on a systematic review. *Psychiatry Research*, *334*, 115816.

*Xian, L., Vickers, S. D., Giordano, A. L., Lee, J., Kim, I. K., & Ramaswamy, L. (2019). "#Selfharm on Instagram: Quantitative analysis and classification of non-suicidal self-injury". *2019 IEEE First International Conference on Cognitive Machine Intelligence (CogMI)*, Los Angeles, CA, USA. doi: 10.1109/CogMI48466.2019.00017.

Xiao, Q., Song, X., Huang, L., Hou, D., & Huang, X. (2022). Global prevalence and characteristics of non-suicidal self-injury between 2010 and 2021 among a non-clinical sample of adolescents: A meta-analysis. *Frontiers in Psychiatry*, *13*, 912441.

*Xu, X. M., Liu, Y. S., Hong, S., Liu, C., Cao, J., Chen, X.R., ... & Kuang, L. (2024). The prediction of self-harm behaviors in young adults with multi-modal data: An XGBoost approach. *Journal of Affective Disorders Reports*, *16*, 100723.

*Yang, J., Chen, Y., Yao, G., Wang, Z., Fu, X., Tian, Y., & Li, Y. (2022). Key factors selection on adolescents with non-suicidal self-injury: A support vector machine based approach. *Frontiers in Public Health*, *10*, 1049069.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122.

*Yin, X., Ma, D., Zhu, K., & Li, D. (2021). Identifying intentional injuries among children and adolescents based on Machine Learning. *PLOS ONE*, *16*(1), e0245437.

*Zhong, Y., He, J., Luo, J., Zhao, J., Cen, Y., Song, ... & Luo, J. (2024). A machine learning algorithm-based model for predicting the risk of non-suicidal self-injury among adolescents in western China: A multicentre cross-sectional study. *Journal of Affective Disorders*, *345*, 369–377.

*Zhou, S. C., Zhou, Z., Tang, Q., Yu, P., Zou, H., Liu, Q., ... & Luo, D. (2024). Prediction of non-suicidal self-injury in adolescents at the family level using regression methods and machine learning. *Journal of Affective Disorders*, *352*, 67–75.

Application of machine learning to improve the predictive performance of non-suicidal self-injury: A systematic review

GAO Baixue[1], XIE Yunlong[1], LUO Junlong[1、2], HE Wen[1、2]

([1] *College of Psychology, Shanghai Normal University, Shanghai* 200234, *China*)

([2] *Lab for Educational Big Data and Policymaking, Ministry of Education, Shanghai Normal University, Shanghai* 200234, *China*)

Abstract: Non-suicidal self-injury (NSSI) is a significant public health problem characterised by widespread stigma, high complexity and heterogeneity. Traditional NSSI research measure and analysis methods are limited, resulting in very low predictive power of the identified factors. In recent years, machine learning has gradually been applied to the analysis and modelling of NSSI. Through simplified questionnaire models and complex multimodal data models, the importance of predictive factors can be visualised and more accurate NSSI classification can be achieved, thus improving the overall predictive performance to a moderate level. In the future, it is necessary to combine traditional NSSI theories and methods to make the screening criteria more stringent, and combine unsupervised learning with transfer learning to increase the reproducibility and comparability of the models. Furthermore, combining non-questionnaire NSSI data with deep learning meanwhile is helpful to improve the predictive performance.

Keywords: machine learning, non-suicidal self-injury, predictive power, application